

# Second order discretization of Backward SDEs\*

D. Crisan, K. Manolarakis<sup>†</sup>

## Abstract

In [5] the authors suggested a new algorithm for the numerical approximation of a BSDE by merging the cubature method with the first order discretization developed by Bouchard and Touzi [3] and Zhang [16]. Though the algorithm presented in [5] compared satisfactorily with other methods it lacked the higher order nature of the cubature method due to the use of the low order discretization. In this paper we introduce a second order discretization of the BSDE in the spirit of higher order implicit-explicit schemes for forward SDEs and predictor corrector methods.

**Key words:** Backward SDEs, Second order discretization, Numerical analysis.

**Mathematics subject classification:** 60H10; 65C30.

## 1 Introduction

The present paper is concerned with the problem of numerical approximation to forward backward SDEs (FBSDEs henceforth). Let  $(\Omega, \mathcal{F}, P)$  be a probability space on which we have defined a triple of processes  $(X, Y, Z)$  which solve the decoupled forward-backward system:

$$\begin{aligned} X_t &= X_0 + \int_0^t V_0(X_u) du + \sum_{i=1}^d \int_0^t V_i(X_u) \circ dW_u^i, \quad t \in [0, T]. \\ Y_t &= \Phi(X_T) + \int_t^T f(X_u, Y_u, Z_u) du - \int_t^T Z_u \cdot dW_u \end{aligned} \quad (1.1)$$

where  $W_t$  is a  $d$ -dimensional Brownian motion and

$$V_k : \mathbb{R}^q \rightarrow \mathbb{R}^q, \quad k = 0, \dots, d, \quad f : \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$$

are some appropriate functions. The system is called decoupled as the (backward) processes  $(Y, Z)$  do not appear in the dynamics of the forward component  $X$ . Systems of the

---

\*The results in this paper were announced by the second author at the workshop "New advances in Backward SDEs for financial engineering applications", Tamerza (Tunisia), October 25 - 28, 2010

<sup>†</sup>Department of Mathematics, Imperial College London, 180 Queen's Gate, London, SW7 2AZ, UK.

form (1.1) have received a lot of attention over the past twenty years primarily due to their applications in the field of Mathematical Finance (see for example [6] and the references therein for details).

Of equal importance is the fact that, the stochastic flow associated with (1.1), i.e., the triple of processes  $(X^{(t,x)}, Y^{(t,x)}, Z^{(t,x)})$ ,  $(t, x) \in [0, T] \times \mathbb{R}$  satisfying

$$\begin{aligned} X_s^{t,x} &= x + \int_t^s V_0(X_u^{t,x}) du + \sum_{i=1}^d \int_t^s V_i(X_u^{t,x}) \circ dW_u^i \\ Y_s^{t,x} &= \Phi(X_T^{t,x}) + \int_s^T f(X_u^{t,x}, Y_u^{t,x}, Z_u^{t,x}) du - \int_s^T Z_u^{t,x} \cdot dW_u \end{aligned} \quad s \in [t, T], \quad (1.2)$$

provides a Feynman-Kac representation for the (viscosity) solution of a class of semi-linear PDEs. In particular, let  $u(t, x) : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , be the viscosity solution of

$$\begin{aligned} \frac{du}{dt} + \tilde{V}(x) \cdot \nabla u + \frac{1}{2} \text{Tr} [V(x) V^*(x) D^2 u] + f(t, x, u, \nabla u V(x)) &= 0, \\ u(T, x) &= \Phi(x) \end{aligned} \quad (1.3)$$

where  $\tilde{V}(x) = V_0(x) - \frac{1}{2} \sum_{j=1}^d \nabla V_j(x) V_j(x)$ . In their seminal work [15], Pardoux and Peng showed that

$$u(t, x) = Y_t^{t,x}, \quad (t, x) \in [0, T] \times \mathbb{R}^d. \quad (1.4)$$

In addition, Ma and Zhang [14] showed that when this viscosity solution is continuously differentiable in its spatial variables, we have the following representation for the solution of the PDE 1.3 and its gradient

$$u(t, x) = Y_t^{t,x}, \quad Z_t^{t,x} = \nabla u(t, x) V(x), \quad a.s. \quad (1.5)$$

From the perspective of numerical analysis, the above means that any probabilistic method for the resolution of (1.1), provides an algorithm for the resolution of semi-linear PDEs. As a result, the interest in robust algorithms for their resolution is high.

Algorithms designed to solve the problem of numerical approximation of the backward part of (1.1) consist of two parts: Firstly, the backward equation is discretized. This step involves the use of one or more conditional expectations. Secondly, a numerical method is grafted onto the chosen discretization to compute the conditional expectations involved. Following the work of Bouchard and Touzi [3] and Zhang [16], the most popular discretization scheme is given by:<sup>1</sup>

$$\begin{aligned} Y_{1,t_n}^\pi &:= \Phi(X_{t_n}), \quad Z_{1,t_n}^\pi := 0 \\ Z_{1,t_i}^\pi &:= \frac{1}{\delta_{i+1}} \mathbb{E} [Y_{1,t_{i+1}}^\pi \Delta W_{i+1} | \mathcal{F}_{t_i}], \quad i = 0, \dots, n-1 \\ Y_{1,t_i}^\pi &:= \mathbb{E} [Y_{1,t_{i+1}}^\pi | \mathcal{F}_{t_i}] + f(t_i, X_{t_i}, Y_{1,t_i}^\pi, Z_{1,t_i}^\pi), \quad i = 0, \dots, n-1, \end{aligned} \quad (1.6)$$

where  $\pi$  is a given partition  $\pi := \{0 = t_0 < t_1 < \dots < t_n = T\}$  of  $[0, T]$  and  $\Delta W_{i+1} = W_{t_{i+1}} - W_{t_i}$ ,  $i = 0, \dots, n-1$ .

---

<sup>1</sup>See Bender and Zhang [2] for an alternative approach based on Picard iterations

Assuming that all coefficients of (1.1) are at least Lipschitz continuous in their spatial variables, we have, following [3] and [16], that

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[ |Y_t^\pi - Y_t|^2 \right] + \mathbb{E} \left[ \int_0^T |Z_t^\pi - Z_t|^2 dt \right] \leq C|\pi|, \quad (1.7)$$

where  $\{(Y_t^\pi, Z_t^\pi), t \geq 0\}$  are the step processes

$$Y_t^\pi := \sum_{i=0}^{n-1} Y_{1,t_i}^\pi 1_{[t_i, t_{i+1})}(t) + Y_{1,t_n}^\pi 1_{t=t_n}, \quad Z_t^\pi := \sum_{i=0}^{n-1} Z_{1,t_i}^\pi 1_{[t_i, t_{i+1})}(t) + Z_{1,t_n}^\pi 1_{t=t_n}. \quad (1.8)$$

In other words, the above discretization of the backward part achieves a convergence of order  $1/n^{1/2}$ , when  $n$  points are used, i.e. the same order as the strong convergence order of the Euler scheme for a (classical) SDE. When more smoothness on the coefficients is assumed, it is shown in Gobet and Labart [8] that the rate of convergence of the processes (1.8) is of order  $1/n$ . In fact an error expansion is obtained in [8] and the leading order coefficients in the error expansion are identified.

To be more precise, the above results are proved for the case when the process  $X$  in (1.6) is replaced by its Euler approximation  $X^\pi$  and, respectively, the filtration  $\{\mathcal{F}_{t_i}\}_{i=0}^n$  is replaced by the natural filtration  $\{\mathcal{F}_{t_i}^\pi\}_{i=0}^n$  associated with  $X^\pi$ . However the same proofs apply both to  $X^\pi$  and  $X$ .

To obtain an implementable scheme using the discretization (1.6), one has to develop a method for approximating the conditional expectations<sup>2</sup>  $\mathbb{E}[Y_{1,t_{i+1}}^\pi | \mathcal{F}_{t_i}]$  and  $\mathbb{E}[Y_{1,t_{i+1}}^\pi \Delta W_{i+1} | \mathcal{F}_{t_i}]$ . Various such methods have been introduced, based on Malliavin calculus [3], on projection on function basis [9], [10] and on quantization [1]. In [5], the authors suggested the application of the cubature method of [13], which is based on the ideas of Kusuoka [12]. The overall rate of convergence of this second approximation step is still of order  $1/n^{1/2}$  or of order  $1/n$  when coefficients are smooth.

In this paper, we introduce a discretization for the backward component of the BSDEs of order  $1/n^2$ . As for the discretization (1.6), this new scheme will require the computation of conditional expectations. This can be done by means of an order 5 cubature formula. Since the order 5 cubature method has local order error  $1/n^3$ , when combined with a second order discretization for BSDEs, will generate a genuine second order algorithm for BSDEs.

To understand which terms such a scheme should include we revisit (1.6) from the point of view of Stratonovich-Taylor expansions. Moreover, to deduce the rate of convergence of the method, we shall rely on the Feynman Kac representation (1.5) and on some estimates on the bounds for the derivatives of (1.3) obtained by Delarue [7].

## 2 Preliminaries

Throughout the paper, we will use the following assumptions:

---

<sup>2</sup>When the process  $X$  in (1.6) is replaced by its Euler approximation, then  $\mathbb{E}[Y_{1,t_{i+1}}^\pi | \mathcal{F}_{t_i}^\pi]$  and  $\mathbb{E}[Y_{1,t_{i+1}}^\pi \Delta W_{i+1} | \mathcal{F}_{t_i}^\pi]$  need to be computed.

- (A) The coefficients of the forward SDE  $V_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $i = 0, 1, \dots, d$  have all entries belonging to  $C_b^\infty(\mathbb{R}^d)$ , the space of bounded infinitely differentiable functions with all partial derivatives bounded. We also assume that the matrix  $V := (V_1 | \dots | V_d)$  is elliptic.
- (B) The driver of the BSDE  $f : [0, 1] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  belongs to  $C_b^{[m/2], m}$ . The exact value for the parameter  $m$  shall be made precise as we proceed.
- (C1) The terminal condition  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is Lipschitz continuous.
- (C2) The terminal condition  $\Phi$  belongs to  $C_b^m(\mathbb{R}^d; \mathbb{R})$ . Again the value of  $m$  shall be determined further on.

We shall denote by  $K$  the constant that bounds all derivatives that appear in our assumptions. When (C1) is in force we shall assume that  $\Phi$  is  $K$ -Lipschitz. Of course under (A), (B) with  $m \geq 1$  and (C1) the system (1.1) has a unique solution such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} (Y_t^2 + |X_t|^2) + \int_0^T |Z_s|^2 ds \right] < \infty$$

To abbreviate notation we shall denote by  $M_i \equiv M_{t_i}$ ,  $i = 0, \dots, n$ , where  $M$  can be any of the processes that appear in this paper. Moreover we write  $\mathbb{E}_s[\cdot]$  for  $\mathbb{E}[\cdot | \mathcal{F}_s]$ . Note that in the current (Markovian) set up  $\mathbb{E}[\cdot | \mathcal{F}_s] = \mathbb{E}[\cdot | X_s]$ . We shall also consider as given a partition  $\pi := \{0 = t_0 < t_1 < \dots < t_n = T\}$  of  $[0, T]$  and consider the related notation

$$\delta_i := t_i - t_{i-1}, \quad \Delta W_i := W_i - W_{i-1}, \quad \Delta W_i^l := W_i^l - W_{i-1}^l, \quad l = 1, \dots, d, \quad i = 1, \dots, n.$$

Lastly, the driver of the BSDE shall be abbreviated as  $\bar{f}(s, x) = f(x, u(s, x), \nabla u V(s, x))$ , where  $u$  is the solution of the semilinear PDE (1.3).

**Remark 2.1.** *Any attempt to discretize the backward part of (1.1) should start with a method to produce (an approximation for) the value of the forward diffusion at the times on the partition  $\pi$ . In this paper we assume that we have at our disposal the forward diffusion  $X$  at the times of the partition  $X_{t_1}, \dots, X_{t_n}$ . This entails no loss of generality, as we aim to combine the second order discretization with the cubature method, which approximates directly the law of  $X$  at the times of the partition.*

*For the benefit of methods that would combine the present discretization with a Monte Carlo simulation, that most probably requires some sort of discretization for the forward process  $X$ , we note that all results regarding the second order scheme (particularly Theorem 3.3 and Corollary 3.4) remain valid, once one replaces  $X$  with a second order approximation. A wealth of such approximations<sup>3</sup> are presented on chapters 12- 15 of [11].*

Working towards a higher order discretization of the backward part of (1.1) we shall rely heavily on the Stratonovich-Taylor expansions. Hence we need to fix notation and present some elementary facts regarding these expansions:

---

<sup>3</sup>Such approximations can be chosen to be *explicit*, in other words completely derivative free.

## 2.1 Stratonovich-Taylor expansions

In the following we will use the set of multi-indices  $\mathcal{A} = \{\emptyset\} \cup \bigcup_{m=1}^{\infty} \{0, 1, \dots, d\}^m$  endowed with the norms  $|\cdot|$  and  $\|\cdot\|$  given by

$$|\beta| = \text{length of } \beta, \quad \|\beta\| := |\beta| + \text{Card}\{j : \beta_j = 0, 1 \leq j \leq |\beta|\}.$$

Clearly  $|\emptyset| = \|\emptyset\| = 0$ . For a  $\beta = (j_1, \dots, j_l) \in \mathcal{A}$  we also write  $\beta_- = (j_1, \dots, j_{l-1})$  and  $-\beta = (j_2, \dots, j_l)$ . We define the subsets of  $\mathcal{A}$ ,

$$\mathcal{A}_m = \{\beta \in \mathcal{A} : \|\beta\| \leq m\} \quad \text{and} \quad \mathcal{A}_m^1 = \{\beta \in \mathcal{A} \setminus \{\emptyset, (0)\} : \|\beta\| \leq m\}.$$

Given two multi-indices  $\alpha = (\alpha_1, \dots, \alpha_k)$  and  $\beta = (\beta_1, \dots, \beta_l)$  we define their concatenation as  $\alpha * \beta = (\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_l)$ . For a suitably chosen function  $f$  and a multi-index  $\beta = (\beta_1, \dots, \beta_l)$ , we define the iterated Stratonovich/Lebesgue integrals as follows

$$J_\beta[f]_{t,s} := \begin{cases} f(s) & |\beta| = 0 \\ \int_t^s J_{\beta_-}[f]_{t,u} du & l \geq 1, j_l = 0 \\ \int_t^s J_{\beta_-}[f]_{t,u} \circ dW^{j_l}(u) & l \geq 1, j_l \neq 0 \end{cases}$$

Let  $L^j$ ,  $j = 0, \dots, d$  be the operators:

$$L^j := \sum_{k=1}^q V_j^k \partial_{x_k}, \quad j = 1, \dots, d, \quad L^0 := \partial_t + \sum_{k=1}^q V_0^k \partial_{x_k}.$$

The iteration of this family of operators is understood as :  $L^\alpha f := L^{\alpha_1} \dots L^{\alpha_n} f$ ,  $\alpha = (\alpha_1, \dots, \alpha_n)$  and we also use the convention  $L^\emptyset f = f$ .

Given a multi index  $\alpha = (\alpha_1, \dots, \alpha_n)$  the following identity is proven in Proposition 5.2.10 of [11]

$$W_t^j J^{\alpha}[1]_{0,t} = \sum_{k=0}^{n+1} J_{0,t}^{(\alpha_1, \dots, \alpha_k, j, \alpha_{k+1}, \alpha_n)}, \quad j = 1, \dots, d. \quad (2.1)$$

A direct consequence of (2.1) (Corollary 5.2.11 of [11]) is that if  $\alpha = (k, k, \dots, k)$ ,  $k = 0, \dots, d$  with  $|\alpha| = m$ , then

$$J^\alpha[1]_{0,t} = \frac{1}{m!} \left( J^{(k)}[1]_{0,t} \right)^k \quad (2.2)$$

The (conditional) expectations of iterated integrals can also be computed. In particular we have a characterization for those iterated integrals that have non zero expectation (see, for example, [4]):

**Lemma 2.2.** *Let  $\alpha = (i_1, \dots, i_r)$  be an arbitrary multi-index with  $\|\alpha\| = m$  and  $t \in [0, T]$ . If  $m$  is odd, then  $\mathbb{E}[J_\alpha[1]_{0,t}] = 0$  and if  $m$  is even then*

$$\mathbb{E}[J_\alpha[1]_{0,t}] = \begin{cases} \frac{t^{m/2}}{2^{r-m/2}(m/2)!} & \text{if } \alpha \in \mathcal{A}_{m,r} \\ 0 & \text{otherwise} \end{cases},$$

where  $\mathcal{A}_{m,r} \subset \mathcal{A}_m$  is the set of multi-indices with  $\alpha = \alpha_1 * \dots * \alpha_{m/2} \in \mathcal{A}_m$ , such that  $\alpha_i = (0)$  or  $\alpha_i = (j, j)$ ,  $j \in \{1, \dots, d\}$ .

A non empty subset  $\mathcal{G} \subseteq \mathcal{A}$  is called a hierarchical set if  $\sup_{\alpha \in \mathcal{G}} |\alpha| < \infty$  and  $-\alpha \in \mathcal{G}$  for any  $\alpha \in \mathcal{G} \setminus \{\emptyset\}$ . Given a hierarchical set  $\mathcal{G}$  we define the remainder set

$$\mathcal{B}(\mathcal{G}) = \{\beta \in \mathcal{A} \setminus \mathcal{G} : -\beta \in \mathcal{G}\}.$$

Note that  $\mathcal{A}_m$  and  $\mathcal{A}_m^1$  are hierarchical sets. If  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is a smooth function with all partial derivatives bounded, then by repeatedly applying the Itô-Stratonovich formula, we obtain the Stratonovich-Taylor expansion

$$\begin{aligned} f(X_t^{0,x}) &= \sum_{\alpha \in \mathcal{K}} L^\alpha f(x) J^\alpha[1]_{0,t} + \sum_{\alpha \in \mathcal{B}(\mathcal{K})} J^\alpha[L^\alpha f(\cdot)]_{0,t} \\ &=: \sum_{\alpha \in \mathcal{K}} L^\alpha f(x) J^\alpha[1]_{0,t} + R_m(t, x, f) \end{aligned} \quad (2.3)$$

for any hierarchical set  $\mathcal{K}$ . The second term on the right hand side of (2.3) is called the  $m$ -th order remainder process and is denoted by  $R_m(t, x, f)$ . Obviously as  $m$  increases, and for small  $t$ , the remainder process gets smaller and smaller.

A hierarchical set that facilitates computations for Stratonovich-Taylor expansions is  $\mathcal{A}(m)$ . For this set we have that

$$\mathcal{B}(\mathcal{A}(m)) = \{(j) \star \beta \mid j = 0, \dots, d, \quad \|\beta\| = m\}.$$

To estimate the remainder that corresponds to  $\mathcal{B}(\mathcal{A}(m))$ , we need to estimate iterated Stratonovich integrals of the form  $J^\alpha[g(\cdot, X)]_{s,t}$  for appropriate function  $g$  such that the previous integral makes sense. For  $0 \leq s \leq t$  with  $t - s < 1$ , we have

$$\left| \sup_{s \leq r \leq t} \mathbb{E}_s \left[ J^\alpha[g(\cdot, X)]_{s,r} \right] \right| \leq C(t-s)^{\|\alpha\|/2} \left( \sup_{s \leq r \leq t} \left\{ \mathbb{E}_s [|L^{\alpha_n} g(r, X_r)| 1_{\alpha_n \neq 0}] + \mathbb{E}_s [|g(r, X_r)| 1_{\alpha_n = 0}] \right\} \right), \quad (2.4)$$

see chapter 5 of Kloeden and Platen [11]. Applying the estimate (2.4) to every index in the set  $\mathcal{B}(\mathcal{A}(m))$ , provides us with the following estimate for the remainder of the Taylor formula,

$$\mathbb{E} [|R_m(t, X, f)|^p]^{1/p} \leq C(t-s)^{(m+1)/2} \max_{m \leq \|\gamma\| \leq m+2} \sup_{0 \leq r \leq t} \mathbb{E} [|L^\gamma f(r, X_r)|^p]^{1/p}. \quad (2.5)$$

To conclude our preliminary results, we shortly discuss the backward PDE (1.3). For the latter, when all coefficients are smooth and bounded, it is well understood that it possesses a smooth and bounded solution. Moreover, a smooth solution  $u(t, x)$  will still exist when working only under **(C1)** for  $t \in [0, T)$ . The solution itself will be bounded on all  $[0, T]$ . Its derivatives however shall not. The next result characterises the behaviour of the derivatives.

**Theorem 2.3** (Delarue [7]). *Let **(A)** and **(B)** hold true and assume further that  $\Phi$  is Lipschitz continuous. Then there exists a unique  $u \in C_b^{[m/2], m}([0, T) \times \mathbb{R}^d)$  that solves (1.3). Moreover, for any  $\alpha \in \mathcal{A}(m)$  we have*

$$\|D^\alpha u(t, \cdot)\|_\infty \leq C \frac{\|\nabla \Phi\|_\infty}{(T-t)^{(\|\alpha\|-1)/2}}, \quad t \in [0, T),$$

for a constant  $C$  independent of  $u$ .

In other words, when the final condition is Lipschitz continuous (and the driver is smooth), the derivatives of the solution of (1.3), explode as  $t \uparrow T$  with the given rate. The exact rates are important to quantify the error of our method.

### 3 One step discretization

To understand the terms that a second order scheme should incorporate, we present first the intuitive arguments that lead to the discretization scheme (1.6). We consider the backward part of (1.1) between two successive times of the partition  $\pi$

$$Y_{t_i} = Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(X_s, Y_s, Z_s) ds - \int_{t_i}^{t_{i+1}} Z_s \cdot dW_s$$

and discretize the Riemann integral using the left hand side point, as in [3], thus leading to an implicit equation for  $Y_{t_i}$  and the stochastic part in the usual way, to obtain

$$Y_{t_i} \simeq Y_{t_{i+1}} + \delta_{i+1} f(X_{t_i}, Y_{t_i}, Z_{t_i}) - Z_{t_i} \cdot \Delta W_{i+1}. \quad (3.1)$$

By conditioning (3.1) with respect to  $\mathcal{F}_{t_i}$  in (3.1) we obtain a first order approximation for  $Y_{t_i}$

$$Y_{t_i} \simeq \mathbb{E} \left[ Y_{t_{i+1}} \middle| \mathcal{F}_{t_i} \right] + \delta_{i+1} f(X_{t_i}, Y_{t_i}, Z_{t_i}), \quad (3.2)$$

but for the presence of  $Z_{t_i}$ . To treat the  $Z_{t_i}$ , we can multiply both sides of (3.1) by  $\Delta W_{i+1}^l$ ,  $l = 1, \dots, d$  and condition with respect to  $\mathcal{F}_{t_i}$ , to obtain

$$Z_{t_i}^l \simeq \mathbb{E} \left[ Y_{t_{i+1}} \frac{\Delta W_{i+1}^l}{\delta_{i+1}} \middle| \mathcal{F}_{t_i} \right], \quad l = 1, \dots, d. \quad (3.3)$$

Observe that scheme (1.6) is just the backward iteration of equations (3.2) and (3.3).

Clearly, to achieve a higher order approximation of the BSDE, we need to add more terms in its (stochastic) expansion ideally without involving any derivatives. As far as the driver is concerned, the next elementary result tells us that the Crank-Nickolson rule indeed achieves a third order local error, thus leading to a second order global error:

**Lemma 3.1.** *Let assumptions (A) and (B) hold true,  $\Phi \in C_{lip}(\mathbb{R}^d)$  and denote by  $\bar{f}(t, x) \equiv f(x, u(t, x), \nabla u(t, x)V(x))$ , where  $u(t, x)$  is the classical solution to PDE (1.3). Then there exists a constant  $C$  independent of the partition and of  $u$  such that*

$$\left| \mathbb{E}_i \left[ \int_{t_i}^{t_{i+1}} \bar{f}(s, X_s) ds - \frac{\delta_{i+1}}{2} \left( \bar{f}(t_i, X_{t_i}) + \bar{f}(t_{i+1}, X_{t_{i+1}}) \right) \right] \right| \leq C \delta_{i+1}^3 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty.$$

*Proof.* Note first that the nonlinear Feynman Kac formula tells us that  $Y_t = u(t, X_t)$  and  $Z_t^l = L^l u(t, X_t)$ ,  $l = 1, \dots, d$ , where the existence of  $u$ ,  $\nabla u$  is guaranteed by Theorem (2.3).

The proof is a simple consequence of the expansion (2.3), using the hierarchical set  $\mathcal{A}(3)$ , applied to the integrand of the integral

$$\begin{aligned} \mathbb{E}_i \left[ \int_{t_i}^{t_{i+1}} \bar{f}(s, X_s) ds \right] = \\ \delta_{i+1} \bar{f}(t_i, X_i) + \mathbb{E}_i \left[ \int_{t_i}^{t_{i+1}} \left( \sum_{\alpha \in \mathcal{A}_0(3)} L^\alpha \bar{f}(t_i, X_i) J^\alpha[1]_{t_i, u} + \sum_{\alpha \in \mathcal{A}_0(5) \setminus \mathcal{A}_0(3)} J^\alpha [L^\alpha \bar{f}(\cdot, X_\cdot)]_{t_i, u} \right) du \right] \end{aligned} \quad (3.4)$$

and to  $\bar{f}(t_{i+1}, X_{i+1})$ ,

$$\begin{aligned} \frac{\delta_{i+1}}{2} (f(X_i, Y_i, Z_i) + \mathbb{E}_i[f(X_{i+1}, Y_{i+1}, Z_{i+1})]) = \\ \delta_{i+1} \bar{f}(t_i, X_i) + \frac{\delta_{i+1}}{2} \mathbb{E}_i \left( \sum_{\alpha \in \mathcal{A}_0(3)} L^\alpha \bar{f}(t_i, X_i) J^\alpha[1]_{t_i, t_{i+1}} + \sum_{\alpha \in \mathcal{A}_0(5) \setminus \mathcal{A}_0(3)} J^\alpha [L^\alpha \bar{f}(\cdot, X_\cdot)]_{t_i, t_{i+1}} \right) \end{aligned} \quad (3.5)$$

If  $\alpha \in \mathcal{A}_0(3)$ , then one of the following holds

$$\alpha = \begin{cases} (i), & i = 0, \dots, d \\ (0, i), (i, 0), & i = 1, \dots, d \\ (i, j), (i, j, k) & i, j, k = 1, \dots, d, \end{cases}$$

Elementary facts about stochastic integrals and Lemma 2.2 tell us that  $\mathbb{E}_s[J^\alpha[1]_{s, u}] = 0$ ,  $u \geq s$  for all  $\alpha$ 's as above except when  $\alpha = (0)$ , or  $(j, j)$ ,  $j = 1, \dots, d$ . For these two cases, we have

$$\mathbb{E}_s[J^{(0)}[1]_{s, u}] = u - s \quad \text{and} \quad \mathbb{E}_s[J^{(j, j)}[1]_{s, u}] = \frac{1}{2} \mathbb{E}[(W_u^j - W_s^j)^2] = \frac{u - s}{2}.$$

Substituting the above into (3.4), (3.5) and taking their difference we have

$$\begin{aligned} \left| \mathbb{E}_i \left[ \int_{t_i}^{t_{i+1}} \bar{f}(s, X_s) ds - \frac{\delta_{i+1}}{2} (\bar{f}(t_i, X_{t_i}) + \bar{f}(t_{i+1}, X_{t_{i+1}})) \right] \right| \\ \leq \left| \mathbb{E}_i \left[ \int_{t_i}^{t_{i+1}} \left( \sum_{\alpha \in \mathcal{A}_0(5) \setminus \mathcal{A}_0(3)} J^\alpha [L^\alpha \bar{f}(\cdot, X_\cdot)]_{t_i, u} \right) du - \frac{\delta_{i+1}}{2} \sum_{\alpha \in \mathcal{A}_0(5) \setminus \mathcal{A}_0(3)} J^\alpha [L^\alpha \bar{f}(\cdot, X_\cdot)]_{t_i, t_{i+1}} \right] \right|. \end{aligned}$$

The estimates on the remainder process complete the proof.  $\square$

The above result combined with a Stratonovich Taylor expansion applied on  $Z$  (as the integrand of the stochastic part) give us a first intuitive approach for the second order

discretization:

$$\begin{aligned}
Y_{t_i} &\simeq Y_{t_{i+1}} + \frac{\delta_{i+1}}{2} \left( f(X_{t_i}, Y_{t_i}, Z_{t_i}) + f(X_{t_{i+1}}, Y_{t_{i+1}}, Z_{t_{i+1}}) \right) \\
&\quad - \sum_{l=1}^d \left\{ L^l u(t_i, X_{t_i}) \Delta W_{i+1}^l + \sum_{k=1}^d L^{(k,l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} J^{(k)}[1]_{t_i, s} dW_s^l \right. \\
&\quad \quad + \sum_{k,j=1}^d L^{(k,j,l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} J^{(k,j)}[1]_{t_i, s} dW_s^l \\
&\quad \quad + L^{(0,l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} s dW_s^l + \sum_{\|\alpha\|=3} L^{\alpha(l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} J^\alpha[1]_{t_i, s} dW_s^l \\
&\quad \quad \left. + \int_{t_i}^{t_{i+1}} R_3(L^l u, t_i, s) dW_s^l \right\}
\end{aligned} \tag{3.6}$$

As in (3.3), we need to innovate a way to recover  $Z_i^l$ ,  $l = 1, \dots, d$ ,  $i = 0, \dots, n-1$ , from (3.6), but this time up to a second order error.

If we multiply both sides of (3.6) by  $\frac{\Delta W_{i+1}^l}{\delta_{i+1}}$  and condition with respect to  $\mathcal{F}_i$ , we shall obtain  $Z_i^l$  but some surviving terms of order  $\delta_{i+1}$  will render the approximation first order. For example consider

$$\frac{1}{\delta_{i+1}} \mathbb{E} \left[ L^{(0,l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} s dW_s^l \Delta W_{i+1}^l \middle| \mathcal{F}_{t_i} \right] = \frac{\delta_{i+1}}{2} L^{(0,l)} u(t_i, X_{t_i}).$$

Hence, we need to find an appropriate weight, that will multiply (3.6) and after conditioning with respect to  $\mathcal{F}_{t_i}$ , provides a second order approximation for  $Z_i^l$  by *canceling out all first order terms*. We make the following judicious choice

$$\mathcal{Z}_i^l := \lambda_1 \frac{\Delta W_{i+1}^l}{\delta_{i+1}} + \lambda_2 \frac{J^{(0,l)}[1]_{t_i, t_{i+1}}}{\delta_{i+1}^2}, \quad l = 1, \dots, d, \quad i = 0, \dots, n-1. \tag{3.7}$$

With a few straightforward computations, we obtain for any  $q = 1, \dots, d$

$$\begin{aligned}
\mathbb{E} \left[ \sum_{l=1}^d L^l u(t_i, X_{t_i}) \Delta W_{i+1}^l \mathcal{Z}_i^q \middle| \mathcal{F}_{t_i} \right] &= \left( \lambda_1 + \frac{\lambda_2}{2} \right) L^q u(t_i, X_{t_i}) \\
\mathbb{E} \left[ \sum_{l=1}^d L^{(0,l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} s dW_s^l \mathcal{Z}_i^q \middle| \mathcal{F}_{t_i} \right] &= \left( \frac{\lambda_1 \delta_{i+1}}{2} + \frac{\lambda_2 \delta_{i+1}}{3} \right) L^{(0,q)} u(t_i, X_{t_i}) \\
\mathbb{E} \left[ \sum_{l=1}^d \sum_{k,j=1}^d L^{(k,j,l)} u(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} J^{(k,j)}[1]_{t_i, s} dW_s^l \mathcal{Z}_i^q \middle| \mathcal{F}_{t_i} \right] &= \left( \frac{\lambda_1 \delta_{i+1}}{4} + \frac{\lambda_2 \delta_{i+1}}{6} \right) \sum_{k=1}^d L^{(k,k,q)} u(t_i, X_{t_i}) \\
\mathbb{E} \left[ \sum_{l=1}^d \left( \sum_{k=1}^d \int_{t_i}^{t_{i+1}} J^{(k)}[1]_{t_i, s} dW_s^l + \sum_{\|\alpha\|=3} \int_{t_i}^{t_{i+1}} J^\alpha[1]_{t_i, s} dW_s^l \right) \mathcal{Z}_i^q \middle| \mathcal{F}_{t_i} \right] &= 0.
\end{aligned}$$

By choosing  $\lambda_1 = 4$ ,  $\lambda_2 = -6$ , we have the following :

**Lemma 3.2.** *Let assumptions (A), (B) hold true and let  $u(t, x)$  denote the classical solution of PDE (1.3). Set*

$$\mathcal{Z}_i^l := 4 \frac{\Delta W_{i+1}^l}{\delta_{i+1}} - 6 \frac{f^{(0,l)}[1]_{t_i, t_{i+1}}}{\delta_{i+1}^2}, \quad l = 1, \dots, d, \quad i = 0, \dots, n-2.$$

Then

$$\left| Z_{t_i}^l - \mathbb{E}_i \left[ \left( Y_{t_{i+1}} + \frac{\delta_{i+1}}{2} f(X_{t_{i+1}}, Y_{t_{i+1}}, Z_{t_{i+1}}) \right) \mathcal{Z}_i^l \right] \right| \leq \delta_{i+1}^2 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty.$$

*Proof.* It should be clear from the discussion preceding Lemma 3.2 combined with the result of Lemma 3.1 that

$$\begin{aligned} & \left| Z_{t_i}^l - \mathbb{E}_i \left[ \left( Y_{t_{i+1}} + \frac{\delta_{i+1}}{2} f(X_{t_{i+1}}, Y_{t_{i+1}}, Z_{t_{i+1}}) \right) \mathcal{Z}_i^l \right] \right| \\ & \leq \left| \mathbb{E}_i \left[ \mathcal{Z}_i^l \left( \int_{t_i}^{t_{i+1}} R_3(\tilde{t}, t_i, s) ds + \int_{t_i}^{t_{i+1}} R_3(L^l u, t_i, s) dW_s^l \right) \right] \right| \end{aligned}$$

The result then follows from the estimate (2.5) on the remainder process and the Cauchy Schwartz inequality.  $\square$

Lemmas 3.1 and 3.2 dictate the following second order algorithm :

•**Initialization**

If (C1) is in force :

$$Y_{2,t_n}^\pi := \Phi(X_n), \quad Z_{2,t_n}^\pi := 0, \quad \text{and} \quad Z_{2,t_{n-1}}^\pi := Z_{1,t_{n-1}}^\pi, \quad Y_{2,t_{n-1}}^\pi := Y_{1,t_{n-1}}^\pi.$$

If (C2) is in force:

$$Y_{2,t_n}^\pi := \Phi(X_n), \quad Z_{2,t_n}^\pi := \nabla \Phi(X_n) V(X_n)$$

•**Backward induction:**

$$\begin{aligned} Z_{2,t_i}^\pi &= \mathbb{E}_i \left[ \left( Y_{2,t_{i+1}}^\pi + \frac{\delta_{i+1}}{2} f(X_{i+1}, Y_{2,t_{i+1}}^\pi, Z_{2,t_{i+1}}^\pi) \right) \mathcal{Z}_i \right], \quad \mathcal{Z}_i := (\mathcal{Z}_i^1, \dots, \mathcal{Z}_i^d)^T \\ Y_{2,t_i}^\pi &= \mathbb{E}_i [Y_{2,t_{i+1}}^\pi] + \frac{\delta_{i+1}}{2} \left( f(X_i, Y_{2,t_i}^\pi, Z_{2,t_i}^\pi) + \mathbb{E}_i [f(X_{i+1}, Y_{2,t_{i+1}}^\pi, Z_{2,t_{i+1}}^\pi)] \right) \end{aligned} \quad (3.8)$$

A small clarification is perhaps in order for the peculiarity of our scheme in the first backward step. If the terminal condition is not smooth, then none of the intuitive arguments that we have presented can apply at time  $t_n = T$ . Hence, when working under (C1), we take the first backward step in an Euler fashion. After this, we expect the PDE to “smooth out” the value function and hence all that we have discussed so far apply. We will see in Corollary 3.4 how a non equidistant partition can compensate for this.

**Theorem 3.3.** *Let assumptions (A),(B) and either of (C1), (C2) hold true. Then there exists a constant  $C > 0$  such that*

$$\begin{aligned} & \max_{0 \leq i \leq n-2} \left[ |Y_{t_i} - Y_{2,t_i}^\pi|^2 + \frac{\delta_{i+1}}{4d} |Z_{t_i} - Z_{2,t_i}^\pi|^2 \right] \\ & \leq C \left( |Y_{t_{n-1}} - Y_{2,t_{n-1}}^\pi|^2 + \frac{\delta_n}{4d} |Z_{t_{n-1}} - Z_{2,t_{n-1}}^\pi|^2 \right) + \sum_{i=1}^{n-1} \delta_i^5 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_i, \cdot)\|_\infty^2 \end{aligned}$$

*Proof.* In the following proof,  $C$  will denote a constant whose value might change from line to line. It will however be independent of the partition and of the bounds of the derivatives of the solution of (1.3). For ease of notation, we set

$$\begin{aligned} \Delta_i^\pi Y &:= Y_{t_i} - Y_{2,t_i}^\pi, \quad \Delta_i^\pi Z := Z_{t_i} - Z_{2,t_i}^\pi, \\ \Delta_i^\pi f &:= f(X_{t_i}, Y_{t_i}, Z_{t_i}) - f(X_{t_i}, Y_{2,t_i}^\pi, Z_{2,t_i}^\pi) \\ \Psi_{i+1} &:= Y_{t_{i+1}} + \frac{\delta_{i+1}}{2} f(X_{t_{i+1}}, Y_{t_{i+1}}, Z_{t_{i+1}}), \quad \Psi_{i+1}^\pi := Y_{2,t_{i+1}}^\pi + \frac{\delta_{i+1}}{2} f(X_{t_{i+1}}, Y_{2,t_{i+1}}^\pi, Z_{2,t_{i+1}}^\pi) \\ \Delta \Psi_{i+1} &= \Psi_{i+1} - \Psi_{i+1}^\pi. \end{aligned}$$

Let us fix a value for  $i = 0, \dots, n-2$ . We consider the difference of the solution of the BSDE at time  $t_i$  and of scheme (3.8):

$$\begin{aligned} \Delta_i^\pi Y &= \mathbb{E}_i[\Delta_{i+1}^\pi Y] + \frac{\delta_{i+1}}{2} \mathbb{E}_i[f(X_i, Y_{2,t_i}^\pi, Z_{2,t_i}^\pi) + f(X_{i+1}, Y_{2,t_{i+1}}^\pi, Z_{2,t_{i+1}}^\pi)] \\ &\quad \mp \frac{\delta_{i+1}}{2} \mathbb{E}_i[\bar{f}(t_i, X_i) + \bar{f}(t_{i+1}, X_{i+1})] - \int_{t_i}^{t_{i+1}} \mathbb{E}_i[\bar{f}(s, X_s)] ds \end{aligned} \quad (3.9)$$

According to the estimates of Lemma 3.1 we have that

$$\begin{aligned} & \left| \mathbb{E}_i[f(X_i, Y_{2,t_i}^\pi, Z_{2,t_i}^\pi) + f(X_{i+1}, Y_{2,t_{i+1}}^\pi, Z_{2,t_{i+1}}^\pi)] - \frac{\delta_{i+1}}{2} \mathbb{E}_i[\bar{f}(t_i, X_i) + \bar{f}(t_{i+1}, X_{i+1})] \right| \\ & \leq C \delta_{i+1}^3 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty \end{aligned} \quad (3.10)$$

Moreover, according to the mean value theorem, we can argue on the existence of a real number and vector  $\mu_1 \in \mathbb{R}$ ,  $\nu_1 \in \mathbb{R}^d$  bounded by  $K$ , such that

$$\frac{\delta_{i+1}}{2} \Delta_i^\pi f = \frac{\delta_{i+1}}{2} (\mu_1 \Delta_i^\pi Y + \nu_1 \cdot \Delta_i^\pi Z) \quad (3.11)$$

Combining (3.9)-(3.11) with Young's inequality with  $\gamma_1 > 0$ , we have

$$\begin{aligned} |\Delta_i^\pi Y|^2 &\leq (1 + \gamma_1 \delta_{i+1}) |\mathbb{E}_i[\Delta_{i+1}^\pi \Psi]|^2 + \left(1 + \frac{1}{\gamma_1 \delta_{i+1}}\right) C \delta_{i+1}^2 (|\Delta_i^\pi Y|^2 + |\Delta_i^\pi Z|^2) \\ &\quad + \left(1 + \frac{1}{\gamma_1 \delta_{i+1}}\right) C \delta_{i+1}^6 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2 \end{aligned} \quad (3.12)$$

Next, observe that for any random variable  $F$  which is measurable with respect to  $\mathcal{F}_{t_{i+1}}$  we have

$$|\mathbb{E}_i[F \mathcal{Z}_i]|^2 = |\mathbb{E}_i[(F - \mathbb{E}_i[F]) \mathcal{Z}_i]|^2 \leq \frac{1}{\delta_{i+1}} (\mathbb{E}_i[F^2] - \mathbb{E}_i[F]^2)$$

Combining this with definition of  $Z_{2,t_i}^\pi$ ,  $i = 0, \dots, n-2$  and the conclusion of Lemma 3.2, we have

$$\delta_{i+1} \mathbb{E} \left[ |\Delta_i^\pi Z|^2 \right] \leq 2d \left( \mathbb{E} \left[ |\Delta \Psi_{i+1}|^2 \right] - \mathbb{E} \left[ |\mathbb{E}_i[\Delta \Psi_{i+1}]|^2 \right] \right) + C\delta_{i+1}^6 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2 \quad (3.13)$$

Putting together (3.12) and (3.13) we get

$$\begin{aligned} & \mathbb{E} \left[ |\Delta_i^\pi Y|^2 + \frac{\delta_{i+1}}{4d} |\Delta_i^\pi Z|^2 \right] \\ & \leq (1 + \gamma_1 \delta_{i+1}) \mathbb{E} \left[ |\mathbb{E}_i[\Delta_{i+1}^\pi \Psi]|^2 \right] \\ & \quad + C\delta_{i+1} \mathbb{E} \left[ |\Delta_i^\pi Y|^2 \right] + \left( \frac{C}{\gamma_1} + \frac{1}{4d} + C\delta_{i+1} \right) \delta_{i+1} \mathbb{E} \left[ |\Delta_i^\pi Z|^2 \right] \\ & \quad + C\delta_{i+1}^5 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2 \\ & \leq (1 + \gamma_1 \delta_{i+1}) \mathbb{E} \left[ |\Delta_{i+1}^\pi \Psi|^2 \right] \\ & \quad + C\delta_{i+1} \mathbb{E} \left[ |\Delta_i^\pi Y|^2 \right] + C\delta_{i+1}^5 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2 \end{aligned} \quad (3.14)$$

where we have chosen  $\gamma_1 = C4d$ .

We can argue once more with the mean value theorem and Young's inequality to deduce that

$$\begin{aligned} & (1 - C\delta_{i+1}) \mathbb{E} \left[ |\Delta_i^\pi Y|^2 + \frac{\delta_{i+1}}{4d} |\Delta_i^\pi Z|^2 \right] \\ & \leq (1 + C'\delta_{i+1}) \left( \mathbb{E} \left[ |\Delta_{i+1}^\pi Y|^2 + \frac{\delta_{i+1}}{4d} |\Delta_{i+1}^\pi Z|^2 \right] \right) + C\delta_{i+1}^5 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2. \end{aligned}$$

for some different constant  $C'$ . By appealing to the discrete version of Gronwall's lemma we complete the proof.  $\square$

Theorem 3.3 justifies the fact that scheme (3.8) produces a second order discretization of the backward component of the BSDE when **(A)**, **(B)** hold true and  $\Phi$  is smooth. Clearly in this case,  $\max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2$  is uniformly bounded, the error on the first backward step is of  $\mathcal{O}(1/n^2)$  when  $n$  points are used on the time discretization and hence, with a uniform partition  $n$  points, we obtain an overall error of  $\mathcal{O}(1/n^2)$ .

The more interesting case occurs when  $\Phi$  is only Lipschitz continuous. Given the results of Theorem 2.3, we expect the bound on the derivatives of  $u$  to explode as  $t \uparrow T$ . To compensate for this and for the fact that our first backward step is of Euler style, we work with a non equidistant partition that becomes more dense as we approach  $T$ . Hence we are still able to achieve a  $1/n^2$  rate of convergence with a  $n$ -points partition.

**Corollary 3.4.** *Let assumptions **(A)**, **(B)** hold true and assume that  $\Phi$  is Lipschitz continuous. Consider the discretization (3.8) along the partition  $\pi$  :*

$$t_i = T \left( 1 - \left( 1 - \frac{i}{n} \right)^\beta \right), \quad i = 0, \dots, n, \quad \beta \geq 5.$$

Then, there exists a constant  $C$  independent of the partition and of the value function  $u$ , such that

$$\max_{0 \leq i \leq n-1} \mathbb{E} \left[ |Y_{t_i} - Y_{2,t_i}^\pi|^2 + |Z_{t_i} - Z_{2,t_i}^\pi|^2 \right]^{1/2} \leq \frac{C}{n^2}.$$

*Proof.* Clearly, we can establish the result by estimating the terms  $\delta_{i+1}^5 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_{i+1}, \cdot)\|_\infty^2$ . According to Theorem 2.3, we have

$$\begin{aligned} \sum_{i=1}^{n-1} \delta_i^5 \max_{\|\alpha\|=4,5} \|L^\alpha u(t_i, \cdot)\|_\infty^2 &\leq \sum_{i=1}^{n-1} \delta_i^5 \frac{C \|\nabla \Phi\|_\infty}{(T - t_i)^4} \\ &= \sum_{i=1}^{n-1} T^5 \left( \int_{1-\frac{i}{n}}^{1-\frac{i-1}{n}} \beta s^{\beta-1} ds \right)^5 \frac{C \|\nabla \Phi\|_\infty}{T^4 \left(1 - \frac{i}{n}\right)^{4\beta}} \\ &\leq \sum_{i=1}^{n-1} \frac{T \beta^5 \left(1 - \frac{i-1}{n}\right)^{5(\beta-1)}}{n^5 \left(1 - \frac{i}{n}\right)^{4\beta}} \leq C/n^4 \end{aligned} \quad (3.15)$$

since  $\beta \geq 5$ .

To complete our proof we estimate the error on the first backward step. Using the standard estimate on the Euler discretization error of Lebesgue integrals we have

$$\begin{aligned} |Y_{n-1} - Y_{1,t_{n-1}}^\pi|^2 &= \left| \mathbb{E}_i \left[ \int_{t_{n-1}}^{t_n} \tilde{f}(s, X_s) ds \right] \mp f(X_{n-1}, Y_{n-1}, Z_{n-1}) \delta_n \right. \\ &\quad \left. - f(X_{n-1}, Y_{1,t_{n-1}}^\pi, Z_{1,t_{n-1}}^\pi) \delta_n \right|^2 \\ &\leq C \delta_n^2 \left( 1 + |Y_{n-1} - Y_{1,t_{n-1}}^\pi|^2 + |Z_{n-1} - Z_{1,t_{n-1}}^\pi|^2 \right). \end{aligned}$$

where once again, we have used the mean value theorem. Rearranging terms, we may argue on the existence of a constant  $C$  such that

$$\begin{aligned} (1 - C \delta_n) \left( |Y_{n-1} - Y_{1,t_{n-1}}^\pi|^2 + \frac{\delta_n}{4d} |Z_{n-1} - Z_{1,t_{n-1}}^\pi|^2 \right) \\ \leq C \delta_n^2 + C \delta_n |Z_{n-1} - Z_{1,t_{n-1}}^\pi|^2 \end{aligned} \quad (3.16)$$

From standard estimates on BSDEs we know that under **(A)**, **(B)** and **(C1)**

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[ |Z_t|^2 \right] < +\infty.$$

Lastly, under **(C1)** we have that

$$\begin{aligned} \mathbb{E} \left[ |Z_{1,t_{n-1}}^\pi|^2 \right] &= \frac{1}{\delta_n^2} \mathbb{E} \left[ |\mathbb{E}_{n-1} [\Phi(X_n) \Delta W_n]|^2 \right] \\ &= \frac{1}{\delta_n^2} \mathbb{E} \left[ |\mathbb{E}_{n-1} [(\Phi(X_n) - \Phi(X_{n-1})) \Delta W_n]|^2 \right] \leq C \end{aligned} \quad (3.17)$$

Substituting (3.17) into (3.16) and then taking square roots, completes the proof.  $\square$

## References

- [1] Vlad Bally and Gilles Pagès. Error analysis of the quantization algorithm for obstacle problems. *Stochastic Processes and their Applications*, pages 1–40, 2003.
- [2] Christian Bender and Jianfeng Zhang. Time discretization and Markovian iteration for coupled FBSDEs. *Annals of Applied Probability*, 18(1):143–177, 2008.
- [3] Bruno Bouchard and Nizar Touzi. Discrete time approximation and Monte Carlo simulation for Backward Stochastic Differential Equations. *Stochastic processes and their applications*, 111:175–206, 2004.
- [4] Dan Crisan and Saadia Ghazali. *On the convergence rate of a general class of SDEs*, pages 221–248. World Scientific Publishing, NJ, 2007.
- [5] Dan Crisan and Konstantinos Manolarakis. Numerical solution for a BSDE using the Cubature method. preprint available at <http://www2.imperial.ac.uk/~dcrisan/>, 2007.
- [6] Dan Crisan and Konstantinos Manolarakis. Probabilistic methods for semilinear partial differential equations. Applications to finance. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44:1107–1133, 2010.
- [7] Francois Delarue. Smoothing property of a non -degenerate semi linear PDE. 2010. preprint.
- [8] Emanuel Gobet and Celine Labart. Error expansion for the discretization of Backward Stochastic Differential Equations. *Stochastic processes and their applications*, 117(7): 803–829, 2007.
- [9] Emanuel Gobet, Jan Phillip Lemor, and Xavier Warin. A regression based Monte Carlo method to solve Backward Stochastic Differential Equations. *Annals of Applied Probability*, 15(3):2172–2002, 2005.
- [10] Emanuel Gobet, Jan Phillip Lemor, and Xavier Warin. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, 12(5):889–916, 2006.
- [11] Peter Kloeden and Eckhard Platen. *Numerical solutions of Stochastic Differential Equations*. Springer, 1999.
- [12] Shigeo Kusuoka. Approximation of expectations of diffusion processes based on Lie algebra and Malliavin calculus. *UTMS*, 34, 2003.
- [13] Terry Lyons and Nicolas Victoir. Cubature on Wiener space. *Proc. Royal Soc. London*, 468:169–198, 2004.
- [14] Jin Ma and Jianfeng Zhang. Representation theorems for Backward Stochastic Differential Equations. *Annals of Applied Probability*, 12(4):1390–1418, 2002.
- [15] Etienne Pardoux and Shige Peng. Backward Stochastic Differential Equations and quasi linear parabolic partial differential equations. *Lecture notes in control and information science*, 176:200–217, 1992.
- [16] Jianfeng Zhang. A numerical scheme for BSDEs. *Annals of Applied Probability*, 14(1): 459–488, 2004.